Teemu Ylänne

3D Camera Tracking for Low-budget Production

Helsinki Metropolia University of Applied Sciences
Bachelor of Engineering
Media Engineering
Bachelor of Engineering Thesis
15 May 2011

Abstract

Author Title Number of Pages Date	Teemu Ylänne 3D camera tracking for low-budget production 36 pages + 2 appendices 15 May 2011
Degree	Bachelor of Engineering
Degree Programme	Media Engineering
Specialisation option	Digital Media
Instructors	Antti Laiho, Lecturer Harri Airaksinen, Principal Lecturer

The purpose of the thesis was to find out if there is a free tool for 3D camera tracking and if such exists, to compare it with a commercial counterpart. Additionally, it was investigated whether the chosen tool is suitable for a low-budget production.

To investigate this, a short film was made where a 3D creature created in 3D modelling software was composited with previously filmed live-action footage. To composite the creature with the footage, the footage was analysed with a 3D camera tracking software. An estimation of camera parameters and its movement were the result of the tracking process. The results were exported to 3D modelling software where a virtual camera for animation was created based on the results. Animation created in the 3D modelling software was exported to a video editing software for the final compositing process.

A suitable 3D camera tracking tool was found for low-budget productions. It was competitive compared to a commercial software. It performed well in a normal desktop computer. There were still many problems with the camera tracking process. Many of them can be avoided with proper preparation before the shooting. Solution to some problems would require a bigger crew which is not always possible with low-budget productions.

Keywords	3D camera tracking, visual effect, short film, animation

Tiivistelmä

Tekijä Otsikko Sivumäärä Aika	Teemu Ylänne 3D-kameranjäljitys pienen budjetin elokuvatuotannossa 36 sivua + 2 liitettä 15.5.2011
Tutkinto	insinööri (AMK)
Koulutusohjelma	mediatekniikka
Suuntautumisvaihtoehto	digitaalinen media
Ohjaajat	tuntiopettaja Antti Laiho yliopettaja Harri Airaksinen

Insinöörityön tavoitteena oli selvittää, löytyykö ilmaisia työkaluja 3D-kameranjäljitykseen, ja jos löytyy, verrata yhtä niistä kaupalliseen tuotteeseen. Lisäksi selvitettiin, onko ilmainen työkalu sopiva pienen budjetin elokuvatuotannon tarpeisiin.

Asian selvittämiseksi valmistettiin lyhytelokuva, jossa valmiiksi kuvattuun materiaaliin liitettiin 3D-mallinnusohjelmalla tehty hahmo. Hahmon liittämiseksi valmiiseen materiaaliin analysoitiin materiaali 3D-kameranjäljitysohjelmalla. Analyysin tuloksena saatiin arvio kameran ominaisuuksista ja liikkeestä. Analyysin tulokset siirrettiin 3D-mallinnusohjelmaan, jossa tulosten perusteella luotiin tiedosto virtuaalikameroineen animaatioita varten. 3D-mallinnusohjelmassa luotu animaatio siirrettiin videoeditointiohjelmaan videokuvaan yhdistämistä varten.

Pienen budjetin elokuvatuotantoihin sopiva ilmainen 3D-kameranjäljitystyökalu löytyi. Vertailussa kaupalliseen vaihtoehtoon se todettiin kilpailukykyiseksi vaihtoehdoksi. Se toimi hyvin tavallisessa tietokoneessa. Kuitenkin itse kameranjäljityksessä kohdattiin useita ongelmia. Monet niistä ovat vältettävissä kunnollisella valmistautumisella ennen varsinaisen materiaalin kuvaamista, mutta osan ratkaiseminen vaatisi työryhmän merkittävää kasvattamista, mikä ei pienen budjetin elokuvatuotannoissa ole aina mahdollista.

Avainsanat	3D-kameranjäljitys, visuaalinen tehoste, lyhytelokuva, animaatio

Contents

Appendices

Abbreviations and Terms

Alpha channel	In digital images a channel where the transparency information is stored.
Feature point	A point in an image that is chosen either manually or by automatic detection to be tracked.
Camera tracking	A process of gathering information from previously filmed footage.
CGI	Computer-generated imagery. Visual effects created with a computer.
Compositing	In cinema a workphase where images from different sources are combined into a single image.
SUSAN	In mathematics an acronym for Smallest Univalue Segment Assimilating Nucleus.
Python	In programming an object oriented programing language.
Visual effects (VFX)	In cinema processes used to manipulate imagery in postproduction process. Includes optical, mechanical and digital effects.

1 Introduction

This thesis has two purposes. Firstly, to find out if there is a non-commercial, free tool for 3D video tracking and if such exists to compare that tool with a commercial counterpart. Secondly, the purpose was to investigate if the chosen tool would be suitable for a low-budget short film project from the perspective of the artist.

The short film, Ihminen ja enkeli, was created to experiment the tracking process on a real video footage. All video was shot without extra lighting and on location. The lighting varies from daylight to fluorescent and tungsten. Most of the shots were taken with a handheld camera, increasing the jitter in the footage. This was a deliberate decision in order to test how the tracking software performs when used with a less than perfect material.

Inspiration for the script of Ihminen ja enkeli comes from an independent graphic short story depicting discussion between Jon Venables, Robert Thompson and a humanoid creature just before the abduction and murder of James Bulger. In the short story the creature plants the seed of violence in the heads of Venables and Thompson through discussion and subtle mindplay. The short story was seen in an exhibition in the mid 1990's and its idea of some creature controlling human mind and actions was haunting enough to be remembered after a decade. The script was written by Teemu Ylänne and Lassi Haaranen, short descriptions of the scenes can be seen in the Appendix 1.

Ihminen ja enkeli was directed by Teemu Ylänne and the camera was operated by Lassi Haaranen. A rough cut of the original material was edited by Teemu Ylänne and used as the source for the camera tracking process. Teemu Ylänne did all the camera tracking. The 3D character was modelled by Lassi Haaranen and animated by Lassi Haaranen and Teemu Ylänne. Music was performed by Lassi Haaranen. Additional sound effects were created by Teemu Ylänne. All editing, mixing and compositing was done by Teemu Ylänne. Budget for the production was 300 euros which was mainly used for props and technical equipment.

The thesis firstly focuses on visual effects in cinema (chapter 2) and proceeds to discuss the process of 3D video tracking and the mathematics behind the process (chapter 3). Then it discusses the use of gathered data in a 3D modelling software (chapter 4) and combining the original video footage with a 3D model and other material (chapter 5). Finally, it focuses on the comparison of the used tracking tools and their features (chapter 6).

2 Visual effects in cinema

2.1 From Méliès to King Kong

Visual effects have been an essential part of the movies for almost their entire history. George Méliès was a French magician whose movie A Trip to the Moon (1902) was one of the first visual effects movies. Méliès used in-camera trickery and special effects to visualise this science fiction tale written by Jules Verne. (Robbins 2010)

Figure 1. Behind the scenes of Metropolis. (Kino International 2010)

Other significant milestones in early visual effects history include Frits Lang's Metropolis (1926) and King Kong (1933), directed by Merian C. Cooper and Ernest B. Schoedsack. Frits Lang's use of mirrors to combine miniature's and actor's in full-sized sets creates a stunning city of the future. As seen in Figure 1, the miniature set included buildings taller than man and model vehicles on suspended roads. In King

Kong the use of miniatures, stop action photography and optical techniques helped the directors tell a story unlike anything ever seen on the silver screen. King Kong set a standard on visual effects that movies are trying to achieve even today. (Robbins 2010; von Bagh 1989: 126-137)

2.2 From Forbidden Planet to Futureworld

In the 1950's visual effects were an essential part of the science fiction movies such as Forbidden Planet (1956) directed by Fred M. Wilcox. Although the movie does not have ground breaking visual effects, it was still a great inspiration for future science fiction entertainment, including Gene Roddenberry's Star Trek. The science fiction films were collaborative works of entertainment and, therefore, the names of the artist are seldom remembered. This changed in 1968 with director Stanley Kubrick's and writer Arthur C. Clarke's collaboration 2001: A Space Odyssey. The film is a visual masterpiece with minimum of dialoque. Kubrick's and Clarke's tremendous visions were fullfilled by Douglas Trumbull. With 205 effects shots that took over half of the movies budget, the visual effects are one of the main reasons why the movie became a classic. (Robbins 2010; von Bagh 1989:594-607; IMDB 2011a)

Michael Crichton's Westworld (1973) introduced audience to 2D computer-generated imagery illustrating a robots point of view. Its sequel Futureworld (1976), directed by Richard T. Heffron, was the first live-action film to use 3D computer-generated imagery. The 3D CGI was used to animate a hand and a face, the CGI face is seen in Figure 2. (Robbins 2010; IMDB 2011b) These movies began the slow transformation from the use of miniatures and special makeup to the modern visual effects done mainly on computers.

Figure 2. CGI from Futureworld (The Lightning Bug's Lair)

Another significant milestone came from the studio that is best known for its animations. In 1982 Disney proved with its classic movie Tron that extensive use of computer generated imagery in a film was possible. Many of its scenes were computer enhanced and some purely CGI. Although the movie looked like a low-budget production compared to its competitors such as Ridley Scott's Blade Runner (1982) that relied in traditional effects, its value is in its role as an antecedent for the advance in visual effects history. (Robbins 2010)

2.3 From Judgement Day to the present

The use of computer in creation of visual effects significantly increased in the 1990's. Advances in the computer technology made creating complex effects easier and more affordable. James Cameron's Terminator 2: Judgement Day (1991) includes a lot of 3D animation and morphing effects. Figure 3 shows the T1000 made of liquid metal morphing into a police officer. In Death Becomes Her (1992) actresses head was removed from the shot and replaced with tracking another shot of a talking head onto it. (Kerlow 2004: 22-24)

Figure 3. T1000 morphing into a police officer.(Movie Mobsters. 2010)

All the techniques to create visual effects for a movie were in use in The Lord of the Rings -trilogy directed by Peter Jackson. In The Return of the King over 1400 visual effects shots were made to create the movie. Pelennor Fields were completely created with computer-generated imagery, because there was no location to be found that would match J.R.R. Tolkien's description of the place. (Visual Effects. Weta Digital 2004)

Figure 4. Spaceships from Star Wreck: In the Pirkinning. (Energia Productions 2011)

The development of the computer technology has made it possible to do visual effects that used to require expensive equipment on an ordinary workstation computer. This has brought the possibility of creating stunning visual effects available for low budget movies but the required software is still expensive. (Byrne 2009, Kerlow 2004) With limited resources it also requires significant amounts of time to do convincing visual effects. A good example of this is the Finnish science fiction parody Star Wreck: In the Pirkinning that was first released on the internet in 2005. The movie took seven years to complete. In that time many of the tools were upgraded and the hardware changed several times. Still everything was done with ordinary desktop computers. A sample of the results, a group of CGI spaceships, is seen in Figure 4. (Energia Productions 2011)

3 Tracking

3.1 Preparation and acquiring the footage

Before shooting the actual video footage thought to be tracked, it is needed to prepare for the shoot as well as possible. In every phase of the preparation one must consider the requirements of tracking as well as everything else that might affect the shooting. In storyboarding it is crucial to visualise all the 3D elements to be included in the final composition. Otherwise the shots may be too tight to fit the intended 3D elements in or the shooting will lack proper purpose. A crude 3D animation of the scene is a good pre-visualisation tool also if there is enough resources to do it. Inadequate planning results in creative deficit. (Byrne 2009: 19-20; Kerlow 2004: 59,62-63)

When acquiring the actual footage it is necessary to keep in mind the limitations of the tracking process. Shots should be made with a locked off camera or a camera on a tripod to avoid unnecessary jitter. The focal length of the camera should be constant if the camera is not on a tripod. A small focal length is preferable. Lighting should be good so that the footage will have enough contrasts. Unnecessary movements near anything that is supposed to be tracked needs to be minimised. When possible one should use suitable markers to help the tracking process. It is important that there is trackable objects both in the background and in the foreground of the shot. The shots should be as short as possible to enhance the possibility of succesful tracking process, preferably less than 400 frames. (Ferguson & Heron; Laboratorium für Informations-technologie 2010; Kerlow 2004:377)

In preparing for the shooting of Ihminen ja enkeli some storyboards were made to clarify the ideas. The 3D character's place in shots was planned to be in relationship with the Ihminen character. In Figure 5 there are three pictures from the crude storyboard made before shooting any material. In the storyboard the main movements of the characters are drawn as arrows pointing the direction of the movement. Only the necessary elements are shown in the storyboard and no consideration for set decoration was made at this stage of production. Some preliminary sketches of the enkeli were also made to visualise the overall tone of the film. There was no real location scouting before the shooting. All locations were chosen based on easy access and proximity. No pre-visualisation animation was made for the short film.

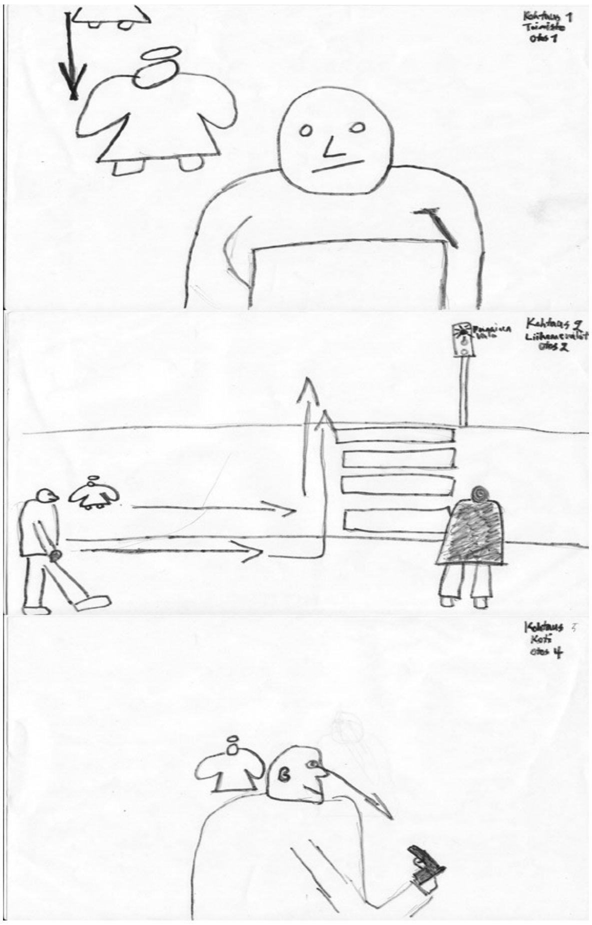

Figure 5. Pictures from crude storyboard

It was decided that shootings would be made with as light equipment as possible. No additional lighting was used and notes of actual lighting conditions were made. The only light source that was excluded from a shot on location was a television because its flicker was disturbing. In one location the lighting was very diverse which made the shooting extra challenging. Because of the nature of the production, it was good that shooting on outdoor locations happened on cloudy weather. This made the lighting on the scenes even which made it easier to light the 3D scenes. Sounds were recorded with the camera's own microphone as a reference for editing. The plan was to discard all the sounds of the actual shots and replace them with sound effects and score. This was decided because there was no way of controlling the sounds of exterior shots.

Most of the shots were taken with a handheld camera although it increased jitter significantly. A tripod was used in some shots to make it easier to controll the movement of the camera. No markers were used to aid the tracking process during the shooting. Before the modelling of the creature or the tracking process started, a rough cut of the material was made. Many of the clips on the rough cut were quite long which was a potential problem for the tracking process but no consideration for shortening these was ever made. For research purposes these shots were perfect because they would really test the chosen tool's suitability for real productions. No color correction or adjustments to the contrast of the shots were made before the tracking process.

3.2 Principles of tracking

Tracking is a process of gathering information from previously filmed footage. It has become one of the key visual effects techniques of the present. Tracking does not produce finished shots but numerical information that can be used to match virtual camera and animated objects and characters with live-action camera. Tracking markers are usually used to help the camera tracking process. (Kerlow 2004: 375, 377) Before starting the tracking process, it is good to preview the shot in real time. Making notes of good tracking targets and when they go out of frame or are occluded can save time and trouble in the actual tracking process. Any parallax in the scene will disqualify otherwise attracting tracking targets. (Wright 2002: 183)

Many commercial tracking tools offer only solution for 2D tracking. 2D tracking tool is in fact a keyframe-generating assistant that analyzes images and calculates pixel shift on defined areas of interest. Visual effects artist could do all this manually but tracking tools are used to speed up the workflow. (Davies 2005: 196) 3D tracking resembles photogrammetry, a technique that can extract three dimensional models from two or more still images of a subject. Usually this is used to extract a depth map.(Kerlow 2004: 383)

In 3D camera tracking suitable points from the image are chosen as feature points either manually or automatically. The position of the feature points in the image plane is measured and combined with the focal length of the camera. 3D tracking software analyses the footage frame by frame and estimates camera parameters from the paths of the chosen feature points. After analysing the acquired footage, the software can produce information usable in creation of a scene in 3D modelling software with virtual camera and the feature points placed in the 3D space. (Laboratorium für Informationstechnologie 2010)

3.3 Automatic feature point detection

Automatic feature point detection chooses points that are bright, dark, edges or corners in the image. Voodoo Camera Tracker uses corner detection at sub-pixel accuracy. Corners are points where edges from at least two different directions meet at the same point. One of the algorithms Voodoo Camera Tracker uses for detection of the feature points is the SUSAN (Smallest Univalue Segment Assimilating Nucleus) corner detector, modification of the edge detector of the same name. (Laboratorium für Informationstechnologie 2010)

In SUSAN corner detector the computation of the area of points inside a circular region $N(x,y)$ have a brightness similar to the one of the central point (x,y). The area is computed thus

$$\eta(x,y)=\sum_{(i,j)\in N(x,y)} \mathrm{e}^{-\left(\frac{I(i,j)-I(x,y)}{t}\right)^{6}} \tag{1}$$

The parameter t controls the sensitivity to noise which means it defines the similarity between the brightness values. The value of $\eta(x,y)$ is compared to a fixed threshold equal to $\eta_{max}/2$ where η_{max} is the maximum value that $\eta()$ can take. That maximum value is

$$c_s(x,y)=\begin{matrix} \frac{\eta_{max}}{2}-\eta(x,y); if\ \eta(x,y)<\frac{\eta_{max}}{2} \\ 0; otherwise \end{matrix} \tag{2}$$

Computational simplicity makes this kind of algorithms based on brightness comparisons preferable solution for corner detection. (Laganière 1998)

To track these points the software needs to estimate their motion in the image plane between two consecutive frames by detecting the changing values of these points. This is done by searching for the highest correlation of image intensity surrounding the points from a rectangular search window. Bad correspondences are detected and excluded from the camera parameter estimation. (Laboratorium für Informations-technologie 2010; Thyagarajan 2006: 256-259)

3.4 Tracking Ihminen ja enkeli

In Ihminen ja enkeli all the clips where the 3D creature was supposed to be were tracked in Voodoo Camera Tracker. Voodoo Camera Tracker is a free 3D tracking tool that can be downloaded from http://www.digilab.uni-hannover.de/download.html. The main window of Voodoo Camera Tracker is shown in Figure 6. Underneath it is the command line window where everything that the software does can be seen as text. The rough cut version of the short film was exported from the video editing sofware as image sequences, one for each shot. These image sequences were loaded to the camera tracking software.

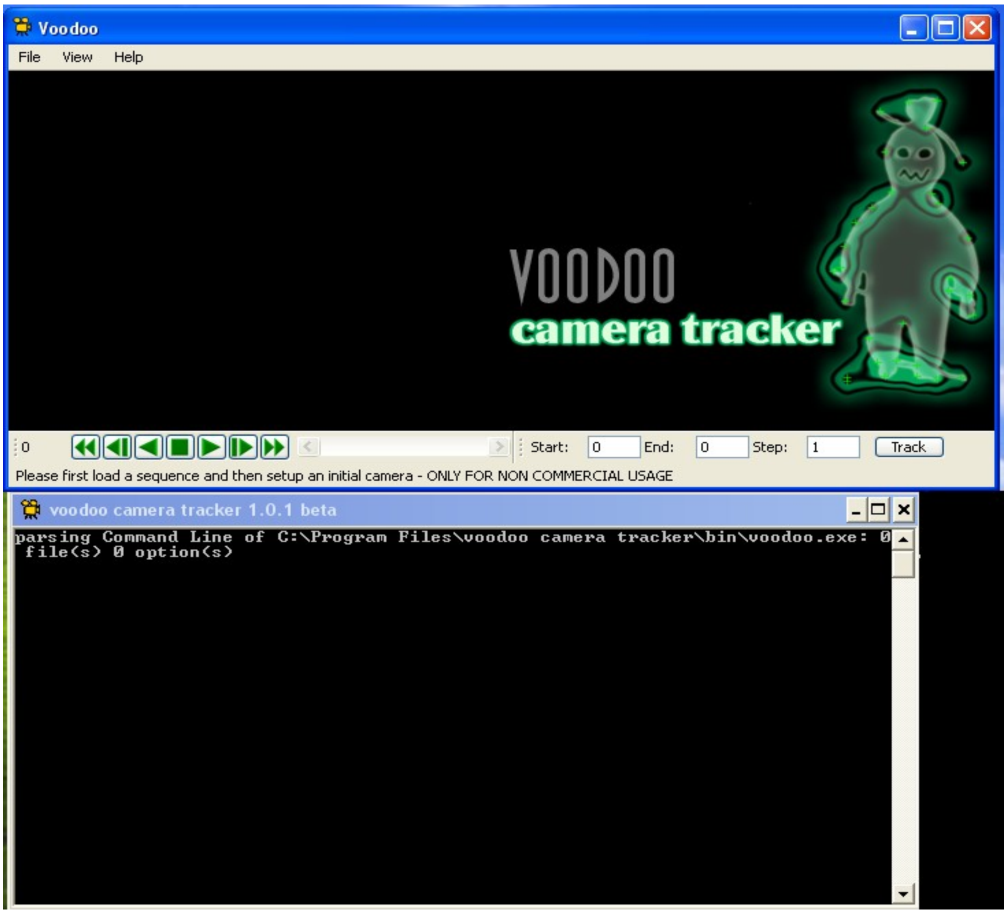

Figure 6. The main window of Voodoo Camera Tracker

After loading a sequence to the software, initial camera settings were made to help the tracking process. The aspect ratio of the footage and the correct color encoding system were chosen from the list provided by the software. Then the tracking process was started and the software analysed feature points from the image using corner detection algorithm. With automatic feature point detection it was possible to use a large amount of feature points. If the feature points had been chosen manually, there would have been a significantly smaller number of tracked feature points resulting in a less accurate tracking data. Figure 7 shows the feature points in an image as crosses. The green crosses are good correspondences also known as inliers. The red crosses are the bad correspondences called outliers. The white crosses are the feature points that have moved from the previous frame.

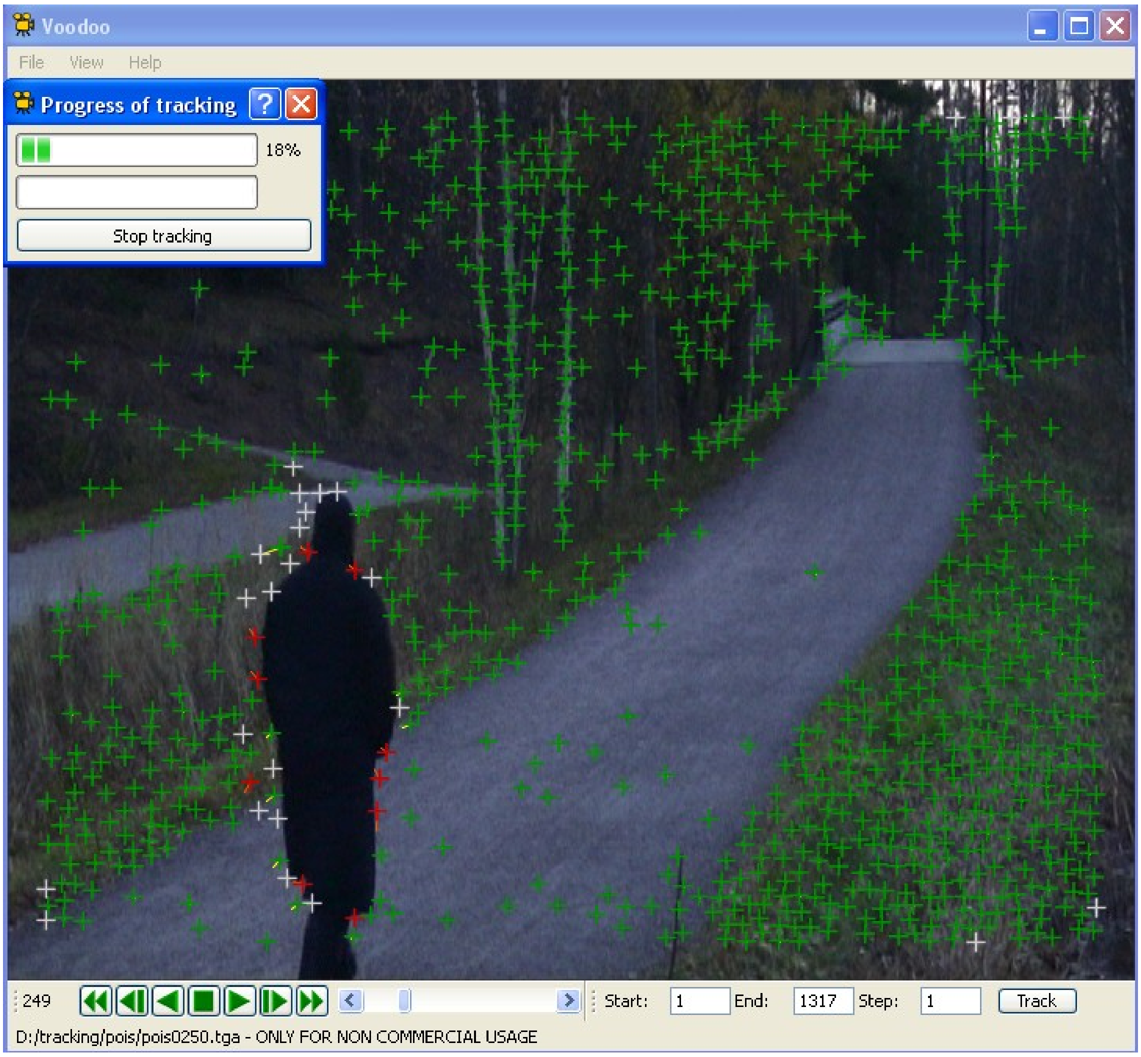

Figure 7. Feature points in an image

After the initial tracking is completed and the camera parameters estimated, Voodoo Camera Tracker allows user to refine the solution by editing feature point tracks and redoing the final estimation step. In Ihminen ja enkeli only some of the shots required editing the feature point tracks. All the parameters and feature points can be saved for future use. After the final estimation the camera parameters and feature points as a 3D scene can be exported for the chosen format. Voodoo Camera Tracker supports exporting to 3D Studio Max, Blender, Lightwave 3D, Maya, Softimage 3D and as a text file that can be used to parse an import filter for any application that Voodoo does not support at the moment. (Laboratorium für Informationstechnologie 2010) In Ihminen ja enkeli, the chosen 3D software suite was Blender, so the tracking results were exported as python scripts. Blender is a free open source 3D modelling suite that can be downloaded from http://www.blender.org/download/get-blender/ (Blender Foundation 2011).

The tracking process took a significant amount of time to complete but most of the results were surprisingly good. Tracking one shot took several hours and most of the time was spent on the actual tracking of feature points. The amount of time used for tracking can be controlled by changing the settings. The user can define, for example, how many corners the algorithm can find and the maximum threshold value. The estimation of the camera parameters and the distribution of the estimation error may need several iterations. Voodoo Camera Tracker has an indistinct way of showing when the tracking process is done. Both progress bars (see Figure 7) reach 100% before the iterations are completed. Fortunately if the user clicks the *Stop tracking* -button the text on the button turns into *Please wait.* When the iterations are completed the whole progress window will close.

For some clips the results of the tracking process were inadequate due to the lack of sufficient lighting. Although this could have been avoided by adjusting the contrast of the shots in the video editing software, it was decided to leave such clips untracked. One whole scene was decided to be removed from the actual storyline because it lacked purpose without succesful tracking data and the creature. A small part of it was used as a short prologue for the film. Many of the shots were much longer than the suggested 400 frame limit. The longest shot was over 1300 frames long but the software could handle it without any problems. Some minor problems with computer's insufficent memory (only 1 Gigabyte of physical memory installed) occured but these were easily eliminated by changing the tracking parameters. A sample of the tracking data produced by Voodoo Camera Tracker can be seen in Appendix 2.

4 Using gathered data in 3D modelling software

After the tracked data is exported from the tracking software, it is brought to the chosen 3D modelling software as a script. From the gathered data a virtual camera is created with the correct focal point and all the movement of the original camera used for acquiring the live-action footage. Also, all the feature points that were tracked are represented in the created 3D scene as a point cloud relative to the camera. Although the feature points are not in the correct position in the 3D scene, they can be used to determine the correct position and movement for the 3D models. (Ferguson & Heron)

It is important to confirm that the 3D scene is in the correct aspect ratio before any animation is done. Easiest way to confirm this is to import the live-action footage as an image sequence to the scene and set it as a background image. After this you can easily see if the point cloud matches the background image from the virtual cameras perspective. When the point cloud matches the background image the 3D model can be imported to the scene. Using the point cloud as a reference, the model can be placed at the correct distance from the virtual camera. (Ferguson & Heron)

Figure 8. Partly masked creature

In Ihminen ja enkeli the data was exported as python scripts to the Blender 3D software suite. The 3D creature, modelled by Lassi Haaranen, was added to the scenes created from the scripts. Additional planes were also added to couple of shots to mask the creature when it was supposed to move through doorways. A partly masked creature is shown in Figure 8. The position of the planes was a trial and error process since the point cloud did not give sufficient information of the correct placement. Although it tells the correct distance from the virtual camera the exact position is always laborious to achieve.

Simple animations were added to make the creature more lively and real. Already in the first proof-of-concept test the 3D creature's belly is wobbling as it hovers in the air. The lighting in the scenes were created to match the live-action footage. Most of the live-action footage has almost colorless, grey look and so the creature became also toned down. For the final shot a beam of light was added for the rise of the creature to exclamate its departure. The beam of light pulls the creature upwards as the Ihminen character continues ahead as seen in Figure 9.

Figure 9. The departure of enkeli

When there is incorrect tracking data, it will cause unwanted movement in the virtual camera. In the second to last shot of Ihminen ja enkeli, there was a clear glitch in the tracking data that caused the virtual camera to jump right. That meant that in the animation the 3D creature jumps to the left. This kind of flaw is easy to correct in the animation phase. On a normal production the correction of such flaws is a standard procedure but since this was a thesis production it was decided to leave the jump in the final animation.

Blender has its own sequence editor which allows compositing 3D models with live-action footage. The feature was tested when making the first proof-of-concept video and it proved to be very competitive solution for compositing. Nevertheless, it was decided not to use this option on Ihminen ja enkeli but to composite the material in video editing software instead. This decision allowed more creative freedom for the editing phase.

The animations were exported as image sequences with matching length of the live-action footage. The background was set as transparent so the only thing that is visible is the desired elements, the 3d creature and in one animation also the light beam. The transparency information was saved in the alpha channel of the images. The exact frames where the 3D creature appears and moves out of the picture in each sequence was documented as a reference for the compositing phase. There were some excess images rendered for most of the shots since the creature is not visible for the whole duration of the shots.

5 Combining the 3D scene and other materials

After the animation in the 3D scene was completed, it needed to be composited with the live-action footage. There are several possibilities for compositing the final images. In big productions this is usually done with a commercial compositing software such as Adobe After Effects or Autodesk's Combustion. But compositing is possible with a wide range of software including 3D software suites, motion graphics software, image manipulation software and video editing software.

The purpose of digital composite is to combine images from variety of different sources in a way that they seem to have been shot at the same time in the same lighting with the same camera (Wright 2002:1). In compositing the foreground and background layers' relative proportions are arbitrated by a matte. With CGI image the matte is called alpha channel. If the lighting is correct in the CGI the compositing is very simple, just placing the CGI on the background image. Colour correcting the CGI at compositing phase needs special approach because it has already been multiplied by its alpha channel. This means that it already is surrounded by black pixels that will result in unpleasant dark edges if multiplied again when composited. This foreground multiplication needs to be turned of if you need to colour correct CGI when compositing. (Wright 2002: 73, 87-88)

In Ihminen ja enkeli, all compositing was done in video editing software. No colour correction was needed in compositing phase, since the lighting of the animations were adjusted in the 3D modelling software to match the lighting conditions in the live-action footage. This allowed the CGI image sequences to be simply placed at the correct place in the timeline. From the documentation made during 3D animation it was easy to check the correct timing of the animations. Animation image sequences were transformed into video clips in order to make them more easily handled. In Figure 10 the live-action clips are seen as light blue rectangles and the animation clips as violet rectangles. The green rectangles represent the sound clips.

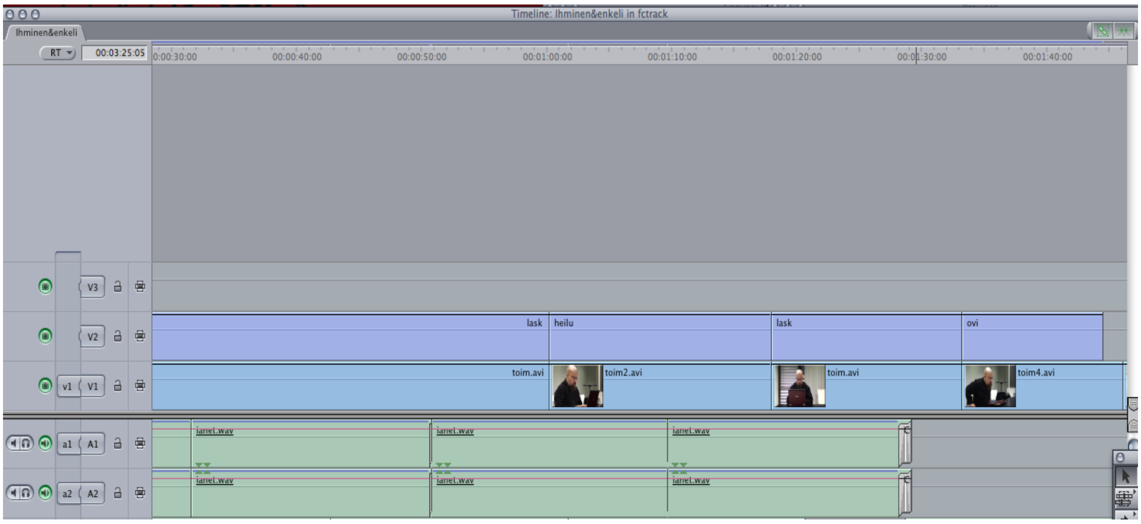

Figure 10. Live-action and animation clips in the time line

To composite the animation and live-action clips the video editing software needs to render each frame. Rendering process in video editing software is always reversible and if a rendered clip is moved on the timeline it needs to be rendered again. This gives possibilty for the editor to manipulate all clips individually anytime needed. In Ihminen ja enkeli this meant that excess animation frames could be removed. It also gave the opportunity to make transitions independently for live-action and animation clips. This was used in the start of the short film to let the animation be unaffected by the fade from black affecting the live-action shot.

The animations matched the live-action footage remarkably well. The masking planes created in the 3D modelling software were almost perfectly placed as can be seen in the composited frame in Figure 11. The part of the creature that is supposed to be behind the opened door is masked with the plane and the difference between the edge of the moving door and the edge of the masking plane is only a couple of pixels. This of course is nearly imposible to detect from the moving picture. There are some continuity issues in the short film concerning the 3D character in the second to last scene. Part of this was caused by inadequate tracking data and part was a deliberate decision in order to include a glimpse of a hand in the shot.

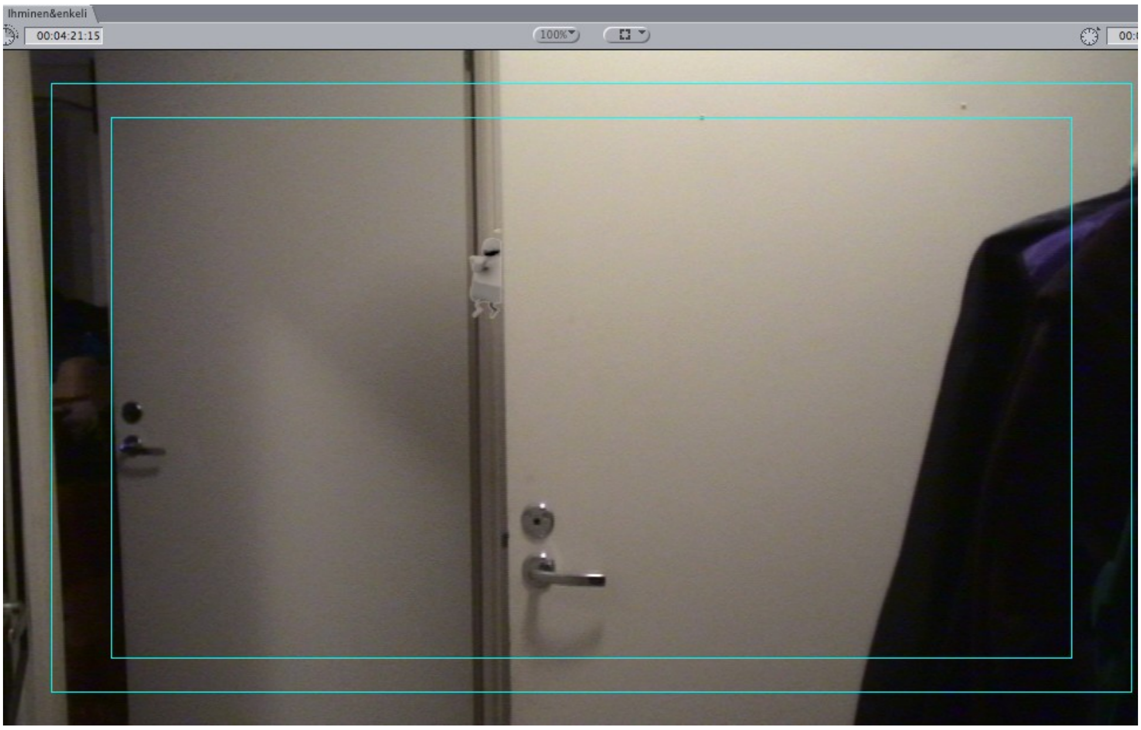

Figure 11. Composited frame from Ihminen ja enkeli

Most of the original sounds were removed and replaced with music and sound effects. Only in the end scene the sound of footsteps is taken from the original footage. This decision was made because there was no possibility to control the sounds during shooting. This also enabled the use of footage soundtrack as collection of audible markers for the editor. This was a great advantage in placing of the sound clips and the animations. The guitar was mixed and edited from an improvised live recording into a score and it was complemented with effects and birdsounds played with a midiplayer. All these were edited together in the video editing software.

6 Comparison

6.1 Voodoo Camera Tracker

Voodoo Camera Tracker is a software developed for research purposes at the Laboratorium für Informationstechnologie, University of Hannover. There is no warranty or support provided, but there is a short manual at their webpage http://www.digilab.uni-hannover.de. The software estimates camera parameters and reconstructs a 3d scene from the footage. It can not handle a real video clip, so all footage must be loaded as image sequences. A typical camera tracking session in Voodoo Camera Tracker is seen in Figure 12. (Laboratorium für Informationstechnologie 2010)

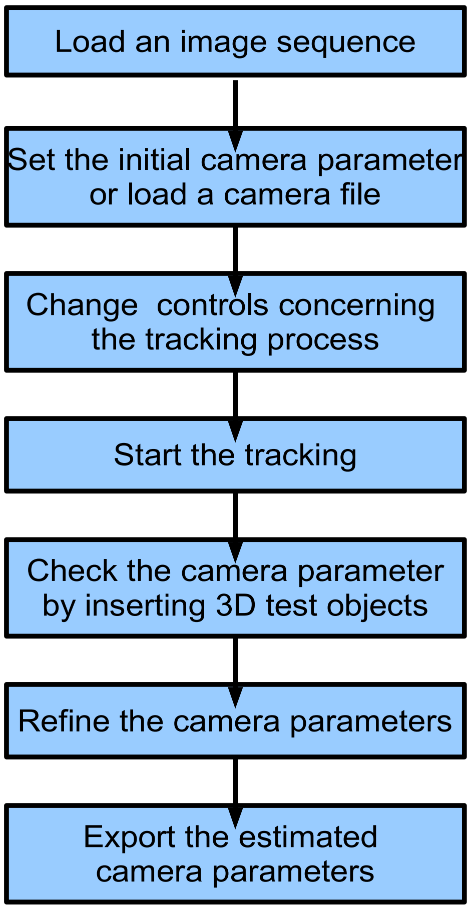

Figure 12. Camera tracking session in Voodoo Camera Tracker (Laboratorium für Informationstechnologie 2010)

Voodoo Camera Tracker detects feature points automatically. It uses a corner detector to determine the feature points. It gives the user a possibility to choose the algorithm used in the detection from five alternatives as seen in Figure 13. Then it compares the image to the previous image and draws path between the matched feature points. The matches are determined by the highest correlation of image intensity in the region surrounding the points. There are also three alternatives for correspondence analysis algorithm. (Laboratorium für Informationstechnologie 2010)

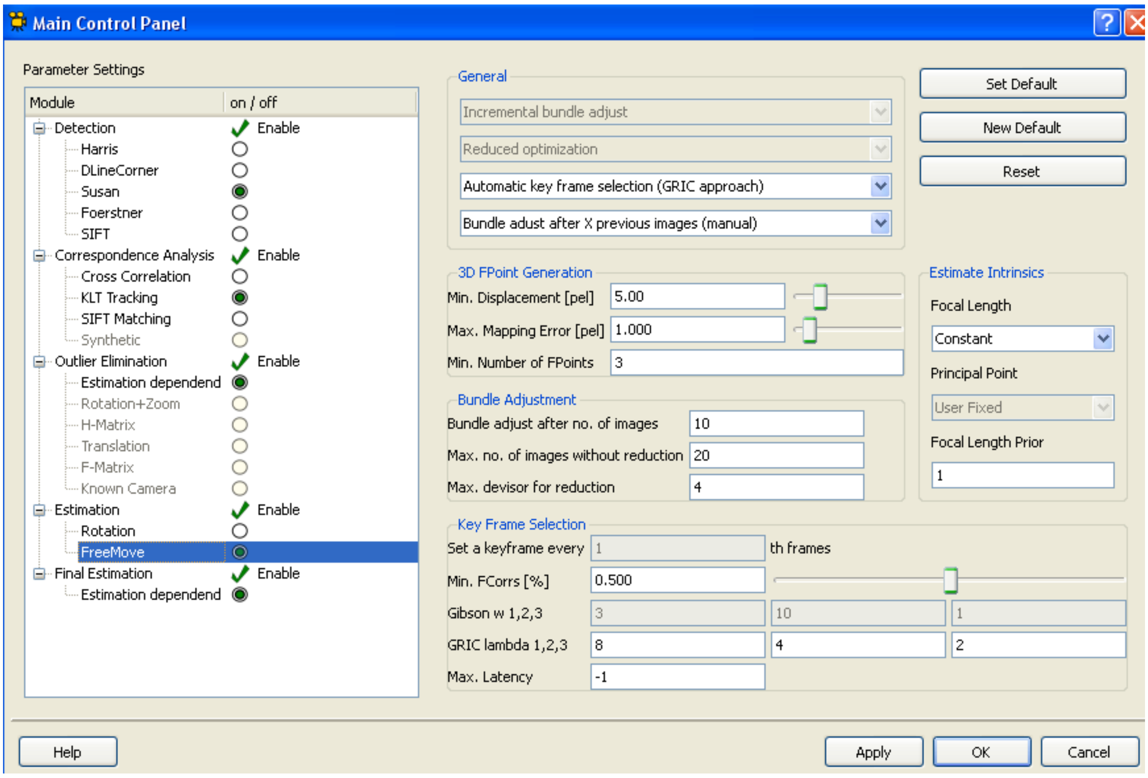

Figure 13. Voodoo Camera Tracker's Main Control Panel

After correspondence analysis Voodoo Camera Tracker tries to detect the bad matches to achieve a robust result. Outlier elimination depends from the chosen estimation alternative, FreeMove or Rotation. Rotation is choosed for footage shot with camera on a tripod, and FreeMove for footage shot with a handheld camera. The camera parameters are estimated in an incremental fashion from the good matches. Finally, it tries to distribute the estimation error of the camera parameters evenly on the whole sequence. (Laboratorium für Informationstechnologie 2010)

3D camera tracking with Voodoo Camera Tracker was very easy due to its automatic feature point detection. The use of the software requires very little information on camera tracking and the online manual gives all necessary information for succesfully completing the tracking process. The possibility to choose different algorithm combinations for feature point detection and correspondence analysis is an interesting feature from an engineer's viewpoint. Voodoo Camera Tracker might not be attracting visually, but its features are very good. As a freeware, it is available for anything but commercial use.

6.2 Autodesk MatchMover

Autodesk MatchMover is a camera tracking tool that is part of the Maya 3D software suite. Autodesk Maya is a 3D animation software with comprehensive tools for animation, modelling, visual effects, matchmoving and compositing. Allthough it costs several thousand euros there is a free 30 day trial available for download. Downloading the free trial from www.autodesk.com requires a registration. (Autodesk, Inc. 2011a) Maya enables the visual effects artist to do almost everything imaginable when creating visual effects for a production. Autodesk has also other tools VFX artist can use to create visual effects, for example Combustion.

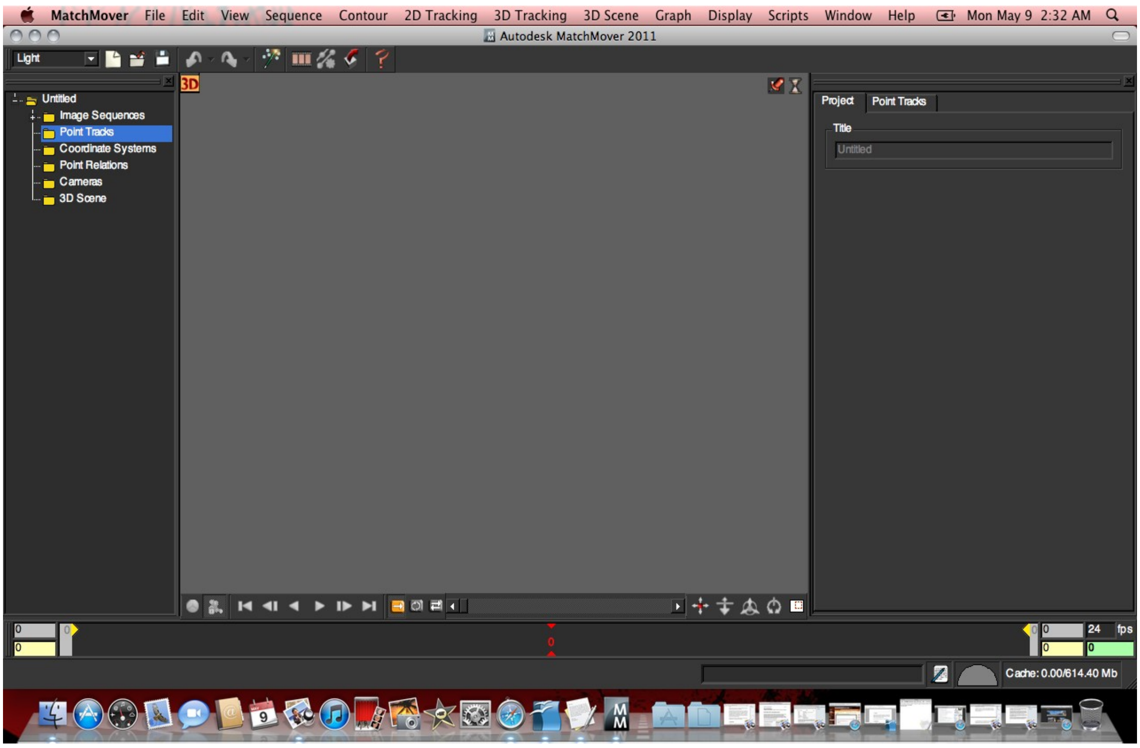

Figure 14. MatchMover

In Autodesk MatchMover there is a possibility to do both 2D and 3D tracking. This is very useful because 2D tracking processes are often needed in making visual effects. The main window of MatchMower is shown in Figure 14. MatchMover can automaticly track in 2D. From the collection of 2D tracks it estimates all camera parameters and 3D point coordinates. Importing is available in many video and image fileformats. MatchMover can export the tracking results as scripts used by many commercial softwares but there is no possibility to export for Blender. (Autodesk, Inc. 2011B)

6.3 Comparison of the chosen tools

The main process of camera tracking is well performed in both tools. Maya's high pricing makes it very difficult to acquire for a low-budget production. The free trial period is a good way to learn oneself the use of the software and if sufficient funding is achieved, Maya is a very attractive solution. There are many good features in MatchMover that Voodoo can not offer at the moment, especially the wider range of importing fileformats. The import window of MatchMover can be seen in Figure 15. Maya has also extensive support resources available and Autodesk's wide variety of 3D animation and effects tools can be bought as a bundle. This kind of comprehensive packaging makes Autodesk's products a perfect choice for productions with adequate funding.

For a low-budget production Voodoo Camera Tracker is the preferred choice for 3D tracking. Voodoo Camera Tracker offers the possibility to choose the algorithm used in feature point detection. This feature is especially interesting for research purposes because it is possible to compare the tracking data produced with different algorithms. Voodoo Camera Tracker also offers very competitive exporting functions. Pricing aside both chosen tools are very well designed and can be used for camera tracking in any production.

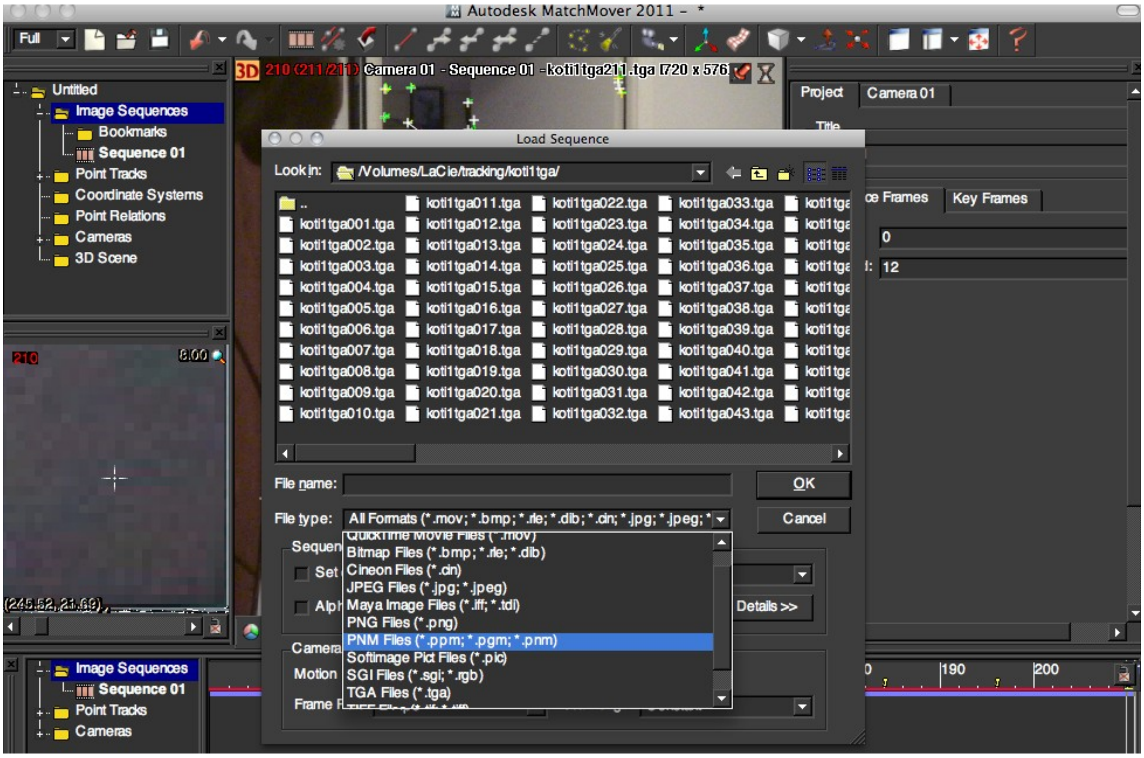

Figure 15. Importing in MatchMover 2011

There is also a commercial camera tracking software based on Voodoo technology called VooCAT by Scenespector Systems which has more features than Voodoo Camera Tracker. Its pricing is suitable for low-budget productions and should be considered as a possible solution for camera tracking needs. (Scenespector Systems 2009) A free trial of VooCAT with disabled export and saving functionality can be downloaded from http://www.scenespector.com/index.php?option=com_content&task=view&id=4&Itemid=33.

7 Conclusions

There is a demand for non-expensive - or even free- visual effects software that would make stunning visual effects available also for low-budget productions and hobbyists. Blender Foundation's 3D software suite Blender is a great open source tool for modelling and animation but some effects need more specialised tools. The process of tracking is a good example of this problem.

There are not that many free software programs one can use for 3D tracking process. There is a discontinued program called Icarus that has been popular in the Blender community (Ferguson & Heron) but after searching for other freeware and different trial software, a decision to use Voodoo Video Tracker on the production of Ihminen ja enkeli was made because it was still being actively developed. 3D modelling suite Maya's tracking tool was chosen as its commercial counterpart because it offered a variety of visual effects tools. Also, with Maya the tracking data could be tested in the same software suite which made the workflow easier.

The use of the chosen free software, Voodoo Video Tracker, was fairly simple and straightforward. It did not require special hardware and performed well on a normal desktop computer. Therefore, it is a suitable solution for a low-budget production for tracking camera parameters from live-action footage and recreating the footage as a 3D scene. The downside is that the use of Voodoo Camera Tracker is limited to non-commercial productions only and it is intended especially for research purposes. This limitation might turn the artist to look for another solution.

For future productions the commercial version of Voodoo Camera Tracker called VooCAT is to be considered as the preferred choice for low-budget productions. Its pricing of under 120 euros (including 19% VAT) and advanced features (Scenespector Systems 2009) makes it a very attractive solution for all camera tracking needs. At least until a new free tool suitable also for commercial use is developed.

There are also many difficulties which must be considered when starting a filming project. The automatic detection of corner points needs high contrasts and is useless on footage with low lighting. Adding tracking points manually does not solve the problem since the automatic correspondence analysis is equally challenged with footage filmed in the dark. The dark shots can be tracked manually but that would require a lot of time and is, therefore, ill-advised. There is always the possibility of animating the virtual camera manually which is much more sensible than manual tracking.

The 3D scene created by the tracking software is relative to the camera and the feature points are represented as a point cloud and are not in correct positions in the 3D space. Every moving object affects the gathered data resulting in outliers and broken tracking paths. This means that the animator can not blindly position the 3D model where the point cloud is. This came apparent when adding masking planes to the animation. Positioning of the planes was a laborious trial and error process but the results were good.

It is impossible to use the same 3D scene for different takes, even if the footage is from the same place and the camera is on a tripod. There are always minor differences between different shots that will result in different results in the tracking data. This means that the decision on which takes will be chosen for the final edit must be made before any tracking is done. A rough edit of the film will also help the person performing the actual tracking process to detect the possible problems beforehand.

The creation of a 3D model to be combined with the video footage would benefit from a clear, trackable placeholder for the model. For example, a simple tennisball on a rod would make creation of the visual effects easier. That would also mean acquiring bigger crew than on this thesis project. Using other tracking markers would make the tracking process more robust and the results more reliable but would also require extra work in post-production removing the markers from the footage. Comprehensive notes from the shootings would also benefit the tracking process but that also requires someone dedicated for the job.

It is possible to use 3D camera tracking in low-budget productions but the smaller the crew and other resources, the more one must rely on the automatic features of the tracking tool. Extensive use of automated features can lead to creative deficit. There must always be enough time to refine the estimated camera parameters when needed.

The best case scenario for the future in the viewpoint of an independent film maker would be that someone would code a tracking tool for Blender 3D software suite. This would allow the independent film maker to create all his/her visual effects in the same software for free. This kind of software tool development would make a suitable final thesis project for engineering students. If developed, it would also be a great asset for educational purposes.

Included with this thesis is a dvd with the completed short film Ihminen ja enkeli, all the python scripts exported from Voodoo Camera Tracker, all the animations and other production materials.

References

Autodesk, Inc. 2011a. Autodesk Maya. [online] URL: usa.autodesk.com/maya. Accessed 8 May 2011.

Autodesk, Inc. 2011b. Autodesk MatchMover online help. Autodesk, Inc.

Blender Foundation. 2011. Blender. [online] URL: http://www.blender.org. Accessed 14 May 2011.

Byrne, Bill. 2009. The visual effects arsenal. VFX solutions for the independent filmmaker. Oxford: Elsevier, Inc.

Davis, Gary M. 2005 Focal easy guide to Combustion 4. For new users and professionals. Oxford: Elsevier, Inc.

Energia Productions. 2011. Star Wreck. [online] URL: http://www.starwreck.com/. Accessed 12 May 2011.

Ferguson, John & Heron Scott. Match moving (camera tracking) tutorial for Blender using Voodoo. [online] URL: http://fictionality.co.uk/doku.php?id= tutorials:matchmoving. Accessed 19 April 2011.

Internet Movie Database (IMDB) 2011a. Forbidden Planet (1956). [online] URL: http://www.imdb.com/title/tt0049223/. Accessed 4 May 2011.

Internet Movie Database (IMDB) 2011b. Futureworld (1976). [online] URL: http://www.imdb.com/title/tt0074559/. Accessed 4 May 2011.

Kerlow, Isaac V. 2004. The art of 3D. Computer animation and effects. Third Edition. Hoboken, New Jersey: John Wiley & Sons, Inc.

Kino International. 2010. Frits Lang's Metropolis -Official Website. [online] URL: http://www.kino.com/metropolis/. Accessed 8 May 2011.

Laboratorium für Informationstechnologie. 2010. Voodoo Camera Tracker: A tool for the integration of virtual and real scenes [online]. Hannover, Germany: University of Hannover. URL: http://www.digilab.uni-hannover.de/docs/manual.html. Accessed 10 April 2011.

Laganière, Robert. 1998. Morphological corner detection. In: 6th international conference on computer vision (1998). [online] URL: http://www.site.uottawa.ca/ research/viva/projects/judoca/coin.pdf. Accessed 8 May 2011.

Movie Mobsters. Throwback Tuesday: Terminator 2 - Judgement Day (1991). 2010. [online] URL: http://www.moviemobsters.com/2010/01/26/throwback-tuesday-terminator-2-judgement-day-1991/. Updated 26 January 2010. Accessed 8 May 2011.

Robbins, David. 2010. Introduction. In Drate, Spencer & Salavetz, Judith, editors. VFX artistry. A visual tour of how the studios create their magic. Oxford: Elsevier, Inc.

Scenespector Systems 2009. VooCAT 1.2. [online] URL:http://www.scenespector.com /index.php?option=com_content&task=view&id= 10&Itemid=25. Accessed 14 May 2011.

The Lightning Bug's Lair. Instant Terror Tuesday: Futureworld (1976). [online] URL: http://www.thelightningbugslair.com/2010/11/instant-terror-tuesday-futureworld-1976.html. Accessed 8 May 2011.

Thyagarajan, K.S. 2006. Digital image processing with application to digital cinema. Oxford: Elsevier, Inc.

Von Bagh, Peter. 1989. Elämää suuremmat elokuvat. Helsinki; Kustannusosakeyhtiö Otava.

Visual Effects. Weta Digital. 2004. [DVD] In The Lord of the Rings. The Return of the King. Special extended DVD edition. The Appendices part 6: The Passing of an Age. New Line Cinema.

Wright, Steve. 2002. Digital compositing for Film and Video. Woburn, MA: Butterworth-Heineman.

List of Scenes

Alkukuva *(toimisto)*: Ihminen istuu työskennellen läppärillä. Enkeli lentää selän takaa ja istuu olkapäälle Ihminen nousee ylös ,lähtee pois (sulkee oven).

Punaisia päin kävely -kohtaus *(Espoossa)*: Kuva takaa, statisti seisoo ja odottaa punaisissa valoissa. Ihminen kävelee punaisia päin, keskellä suojatietä Enkeli kääntyy ympäri ja vilkuttaa siivellä statistille.

Ilkivalta (joku tyylikäs ikkuna): Näkee ikkunan, poimii kiven.

Tilaisuus tekee varkaan -kohtaus *(joku katu/tie)*: Ihminen löytää lompakon kadulta, katsoo sisään, ottaa rahat, laittaa povariin, heittää lompakon pois.

Tokavika kohtaus *(Teemulla kotona)*: Enkeli suputtaa korvaan. Ihminen kaivaa jostain aseen (9 mm) ja katselee sitä. Nousee ylös ja laittaa taskuun.

Loppukuva *(Lepuskissa urheilukentän takana)*: Ihminen kävelee (määrätietoinen) kamerasta pois päin ruskaista hiekkatietä ja Enkeli tanssii olalla .

Sample of Tracking Data

```
# Text export
# created by voodoo camera tracker - www.digilab.uni-hannover.de
# Creation date: Wed Mar 11 14:45:46 2009
# The camera (one line per frame)
#

# Cx Cy Cz Ax Ay Az Hx Hy Hz Vx Vy Vz K3 K5 sx sy Width Height ppx ppy f fov H0x H0y H0z V0x
V0y V0z
#
#timeindex 1
0.0000000000 0.0000000000 0.0000000000 0.0000000000 0.0000000000 1.0000000000
    4176.3782534547 0.0000000000 0.0000000000 0.0000000000
    4454.8173916125 0.0000000000 0.0000000000 0.0000000000 0.0106667000
    0.0100000000 360 288 0.0000000000 0.0000000000 44.5481739161
    4.9357893069 1.0000000000 0.0000000000 0.0000000000 0.0000000000
    1.0000000000 0.0000000000
#timeindex 2
0.0000000000 0.0000000000 0.0000000000 0.0000114125 0.0000085888 0.9999999999
    4177.0289498708 0.8980434511 -0.0476778776 -0.9579164447
    4455.5114700846 -0.0382565958 0.0000000000 0.0000000000 0.0106667000
    0.0100000000 360 288 0.0000000000 0.0000000000 44.5551157322
    4.9350212472 0.9999999768 0.0002149957 -0.0000114143 -0.0002149958
    0.9999999769 -0.0000085864

------------------------------------------------------------------------------------------------------------------------------
----# 3D Feature Points
# X Y Z
15.5080464919 8.1436974188 445.1372379531
16.0425438639 3.2478055207 445.1809390257
1.6531746683 6.0494004968 445.4375957340
4.7459888165 5.9491317161 445.4167299823
5.4920804035 7.8481850317 445.3787410403
3.8931938525 -0.5499785773 445.4643874557
```